

A MODEL FOR QUADRAT SAMPLING WITH "VISIBILITY BIAS"

by

R. Dennis Cook and Frank B. Martin^{*}

Technical Report No. 191

December 1972

University of Minnesota
Minneapolis, Minnesota

^{*}Dennis Cook is an Assistant Professor in the Department of Applied Statistics and Frank Martin is an Assistant Professor and Director of the Statistical Center at the University of Minnesota. The authors wish to thank Professor James Peek and the Department of Entomology, Fisheries and Wildlife for presenting the problem and making possible a remarkable opportunity to study the model at first hand above a selected Minnesota moose range.

SUMMARY

In aerial census data "visibility bias" is present because of the failure to observe some animals. A model is presented for quadrat sampling of randomly occurring groups whose size follows a single parameter power series distribution when there is a probability $q > 0$ of missing single animals. Maximum likelihood estimates of group density, average group size, and the visibility bias parameter, q , are presented. An example with moose census data is worked out.

1. Introduction.

"Visibility bias" is a recognized problem in aerial census techniques for estimating wildlife population density. This source of error is generally conceded to be the main cause of inaccurate aerial census data and, depending on various contributing environmental factors (e.g., snow cover, weather, observer fatigue and terrain) can produce severely biased population density estimates. From among the aerial census procedures usually employed, Evans, Troyer and Lensink [2] judged quadrat sampling to be preferable because of its inherently smaller visibility bias. This bias is, however, still substantial and reduction to a negligible magnitude by an intense sampling effort can be prohibitively expensive (see, for example, Bergerud and Manuel [1]). Also, procedures for determining the size of the visibility bias are, by and large, ad hoc and can require impractical amounts of corroborating ground data.

The problem of visibility bias is confounded if, as is common, animals tend to occur in groups. When, in addition, the quadrat sampling method is employed, estimates of the mean number of groups per quadrat and the mean group size are usually desired. Typically, visibility bias will result in underestimates of the mean number of groups per quadrat because some groups will not be seen during the census. On the other hand, because larger groups should have a higher probability of being observed than should smaller groups, visibility bias will result in overestimates of the mean group size.

We propose here a refinement of the quadrat sampling method which allows a coherent theory of estimation in the face of visibility bias. The model developed requires that animals occur in groups and should improve usefulness or accuracy of aerial census surveys which must be conducted when observer abilities or environmental conditions are somewhat less than ideal.

2. The Model.

The sampling model is based on three main assumptions: First, it is assumed that animals occur within quadrats in groups of varying size. Second, prior to observing any animal within a specified group, it is assumed that each animal within that group has a constant probability, p , of being observed. Third, it is assumed that, conditional on observing at least one member of a group the entire group is observed with certainty. This third assumption implies a change in the sampling procedure conditional on seeing at least one animal. The change entails careful enumeration of all objects in the group by close examination after which the observer returns to the original sampling procedure. This is in contrast to the usual quadrat sampling method in which the observer does not change the sampling procedure and, thus, may fail to observe a part of any sighted group. The proposed sampling method is a refinement of the usual method to the extent that we assume that if at least one member of a group is observed then the entire group is observed with certainty.

The main implication of the above three assumptions is that larger groups will have a higher probability of being observed than will smaller groups. In fact, to fail to observe a group of size S each member of the group must be individually and independently missed with probability $q = 1 - p$. It follows that the probability of observing a group of size S is simply $1 - q^S$. Hence, an individual animal's "visibility" is a function of p and the size S of group it happens to be in.

In addition to these main assumptions a complete specification of this model requires i) the distribution of the number of groups per quadrat, ii) the distribution in the population of group sizes, and iii) the assumption that the number of groups present in a quadrat is independent of the group

sizes present. The distribution of the number of groups per quadrat is assumed to be Poisson (parameter λ), the validity of which is basically a discussion of quadrat size.

The distribution of group size in the population will depend on the species being observed and can vary with time of year and environmental stresses. Thus, it is difficult to specify a single distribution for group size that may be appropriate in a wide variety of circumstances. In this treatment, we will consider only single parameter power series distributions. Specifically, since all groups must have at least one member we assume that the distribution of $e = S - 1$ (excess required to specify a group) is

$$\Pr(e = n | \theta) = \frac{b(n)\theta^n}{g(\theta)} \quad n = 0, 1, 2, \dots \quad (2.1)$$

Some economy in formulation is gained by using this approach. The reader should keep in mind during what follows that the average size in the population of groups is $E(e | \theta) + 1$.

Because of visibility bias ($p < 1$) the observer does not see a simple random sample from either the Poisson distribution of the number of groups per quadrat or the power series distribution of group excess. In fact, it can be easily shown that, regardless of the choice of the group size distribution, the distribution of the observed number of groups per quadrat is again Poisson with parameter expressible in the form

$$\hat{\lambda} = \lambda(1 - qG(q)) \quad (2.2)$$

where $G(q)$ is the probability generating function of group excess, e , evaluated at q . The quantity $(1 - qG(q))$ appearing in (2.2) can be interpreted as being the unconditional probability of observing a group. Notice that, since $\lambda' < \lambda$, the observer will be, on the average, seeing too few groups

per quadrat. It is worth pointing out here that, under this model, the distribution of the observed number of groups per quadrat will be Poisson if and only if the distribution of the actual number of groups per quadrat is Poisson. Practically, this means that the assumption of a Poisson distribution can be checked by using observed data.

The distribution of observed excess (o.e.), under this model, is found to be

$$\Pr(\text{o.e.} = n | \theta, q) = \frac{b(n)\theta^n(1-q)^{n+1}}{g(\theta)[1-qG(q)]} \quad n = 0, 1, 2, \dots \quad (2.3)$$

$$= \Pr(e = n | \theta) \frac{(1-q)^{n+1}}{[1-qG(q)]} \quad (2.4)$$

where $G(q) = g(\theta q)/g(\theta)$. Expression (2.4) makes manifest the effect of visibility bias on the distribution of observed excess--the distribution of o.e. is obtained by systematically shifting density to the right. If $q = 0$ (no visibility bias) the distribution of o.e. reduces to the distribution of e . We note for later use that if q is assumed to be known then o.e. follows a single parameter power series distribution.

The probability generating function of the distribution (2.3) is

$$f(t) = \frac{g(t\theta) - qg(t\theta q)}{g(\theta) - qg(\theta q)} \quad (2.5)$$

and the k^{th} factorial moment is found to be

$$\mu(k) = \frac{\theta^k [g^{(k)}(\theta) - q^{k+1} g^{(k)}(\theta q)]}{g(\theta) - qg(\theta q)} \quad (2.6)$$

The mean and variance can, of course, be easily determined using the first two factorial moments.

3. Estimation.

In this section we consider the maximum likelihood (ML) estimates of θ , p and λ .

Let

m = the number of quadrats sampled.

k_i = the number of groups observed in the i^{th} quadrat, $i = 1, 2, \dots, m$.

$k = \sum k_i$.

n_i = the number of groups of excess i , i.e., the number of groups of size $i + 1$, $i = 0, 1, 2, \dots, r$.

$\alpha_i = n_i/k$.

$\bar{n} = \sum i\alpha_i$.

The number of groups, k , observed in any census is, of course, a Poisson random variable (parameter λ' defined in (2.2)). However, since k is independent of o.e.; θ , p and λ may be estimated using the following procedure: θ and $q (= 1 - p)$ will be estimated simultaneously from the conditional distribution of o.e. given k . Then, given the estimates $\hat{\theta}$ and \hat{q} , λ can be easily estimated using $\bar{k} = \lambda'$ to produce

$$\begin{aligned}\hat{\lambda} &= \bar{k}/(1 - \hat{q}G(\hat{q})) \\ &= \bar{k}g(\hat{\theta})/[g(\hat{\theta}) - \hat{q}g(\hat{\theta}\hat{q})]\end{aligned}\tag{3.1}$$

where $\bar{k} = k/m$. This is equivalent to finding the joint maximum likelihood estimates of θ , p and λ . The estimation problem reduces to finding the joint maximum likelihood estimates of θ and p from the distribution of o.e. The likelihood function can be written in the form

$$L(n_i | \theta, q) = \frac{C_{\theta} \bar{k}^{\bar{n}} \prod_{i=0}^r (1 - q^{i+1})^{n_i}}{[g(\theta) - qg(\theta q)]^k}\tag{3.2}$$

where C is independent of θ and q . The set (n_0, n_1, \dots, n_r) is, of course, always sufficient. From (3.2), however, it is seen that for this problem (n_0, n_1, \dots, n_r) is not only sufficient but minimal sufficient.

In passing, it is interesting to note that if p is assumed to be known, then Σn_1 is minimal sufficient and, recalling that o.e. follows a power series distribution, a minimum variance unbiased estimator of θ can be obtained [3]. Since it is doubtful that p will ever be known we make no further mention of this approach.

The maximum likelihood estimators satisfy the equations

$$\bar{n} = \frac{\hat{\theta}[g^{(1)}(\hat{\theta}) - \hat{q}^2 g^{(1)}(\hat{\theta}\hat{q})]}{g(\hat{\theta}) - \hat{q}g(\hat{\theta}\hat{q})} \quad (3.3)$$

$$\frac{\hat{\theta}\hat{q}g^{(1)}(\hat{\theta}\hat{q}) + g(\hat{\theta}\hat{q})}{g(\hat{\theta}) - \hat{q}g(\hat{\theta}\hat{q})} = \sum_{i=0}^r \frac{\alpha_i(i+1)\hat{q}^i}{1 - \hat{q}^{i+1}}. \quad (3.4)$$

The R.H.S. of equation (3.3) is $E(o.e. | \hat{\theta}, \hat{q})$. After some manipulation equation (3.4) can be written more conveniently as

$$\hat{\theta} \frac{g^{(1)}(\hat{\theta})}{g(\hat{\theta})} + 1 = [1 - \hat{q} \frac{g(\hat{\theta}\hat{q})}{g(\hat{\theta})}] \sum_{i=0}^r \frac{\alpha_i(i+1)}{1 - \hat{q}^{i+1}}. \quad (3.5)$$

Notice now that the L.H.S. of (3.5) is $E(e | \hat{\theta}) + 1$ -- the maximum likelihood estimate of average group size. The R.H.S. of (3.5) can be written in the form $\Sigma(i+1)\beta_i$, where β_i is an estimate of $\Pr(e = i | \theta)$, (see equation (2.4)). Either equation (3.4) or (3.5) may be used in combination with equation (3.3) to find the ML estimates. However, the authors feel that (3.5) is the more useful.

Generally, it will be necessary to use an iterative procedure to find $\hat{\theta}$ and \hat{q} . It is easily seen that

$$E(e | \hat{\theta}) = \frac{\hat{\theta}g^{(1)}(\hat{\theta})}{g(\hat{\theta})} < \bar{n}$$

where $\hat{\theta}$ is the ML estimate of θ . Because of this relation it is recommended that an initial value for $\hat{\theta}$ (say $\hat{\theta}_1$) be chosen so that $E(e | \hat{\theta}_1)$ is close to but less than \bar{n} . With this condition any intermediate initial value for \hat{q} should suffice.

The large sample covariance matrix, $\frac{1}{k} V$, of $(\hat{\theta}, \hat{p})$ is given by the inverse of the information matrix, $kU = k(u_{ij})$:

$$u_{11} = \sigma^2/\theta^2 \quad (3.6)$$

$$u_{12} = u_{21} = \frac{A}{\theta} [\gamma_2(\theta q) + 2\gamma_1(\theta q) - \mu(1+\gamma_1(\theta q))]$$

$$u_{22} = \frac{A\gamma_1(\theta q)}{q} - A^2(1+\gamma_1(\theta q))^2 + \sum_{i=0}^{\infty} q^i g(\theta q^{i+2}) [1+\gamma_2(\theta q^{i+2}) + 3\gamma_1(\theta q^{i+2})]$$

where μ and σ^2 are the mean and variance of the distribution of o.e.; $\gamma_j(x)$ is the j^{th} factorial moment of the distribution of e with parameter x ; and $A = g(\theta q)/[g(\theta) - qg(\theta q)]$. The series occurring in u_{22} should converge very quickly so that, in general, this may be evaluated using a few leading terms.

4. Example.

In this section we present an example and consider in somewhat greater detail estimation of θ , p and λ when the distribution of e is assumed to be Poisson with parameter θ . The data in Table 1 is part of that collected by an aerial census of moose in Superior National Forest, 1969, and was provided by James Peek, University of Minnesota.

TABLE 1: Aerial Census Data Moose From Superior National Forest.

Group Excess (i)	0	1	2	3	4	5	6
Frequency (n_i)	15	17	6	2	0	1	1
Relative Frequency (α_i)	.357	.405	.143	.048	0	.024	.024

$m = 30$, $k = 42$, $\bar{k} = 1.4$, $\bar{n} = 1.095$.

By assuming a Poisson distribution for group excess, $g(\theta) = e^\theta$ and $b(n) = (n!)^{-1}$, the ML estimators of θ and $q(= 1 - p)$ satisfy the equations

$$\bar{n} = \frac{\hat{\theta}[1 - \hat{q}^2 e^{-\hat{\theta}\hat{p}}]}{1 - \hat{q} e^{-\hat{\theta}\hat{p}}} \quad (4.1)$$

$$\hat{\theta} + 1 = [1 - \hat{q}e^{-\hat{\theta}\hat{p}}] \sum_{i=0}^{\infty} \frac{\alpha_i(i+1)}{1 - \hat{q}^{i+1}} . \quad (4.2)$$

The ML estimate $\hat{\lambda}$ of λ is

$$\begin{aligned} \hat{\lambda} &= \bar{k} / (1 - \hat{q}e^{-\hat{\theta}\hat{p}}) \\ &= \bar{k}[\bar{n} - \hat{q}\hat{\theta}] / \hat{\theta}\hat{p} . \end{aligned} \quad (4.3)$$

A first approximation to the variance of $\hat{\lambda}$ can be found by writing

$$\sigma_{\hat{\lambda}}^2 = E \text{Var}(\hat{\lambda}|k) + \text{Var}[E\hat{\lambda}|k]$$

and then using the method of statistical differentials to approximate each term in the R.H.S. The resulting approximation is

$$\sigma_{\hat{\lambda}}^2 \doteq \frac{\lambda}{m(1 - qe^{-\theta p})^3} [q^2 p^2 v_{11} + (1 + q\theta)^2 v_{22} + 2pq(1 + \theta q)v_{12} + (1 - qe^{-\theta p})^2] \quad (4.4)$$

where the v_{ij} come from the covariance matrix, V , defined in expression (3.6). Notice that if θ and p are regarded as known then, upon placing $v_{ij} = 0$, the above approximation reduces to the exact variance.

The solution to equations (4.1) and (4.2) requires an iterative procedure. Initial values $(\hat{\theta}_1, \hat{q}_1)$ for any iterative procedure should be chosen so that

$$\frac{\bar{n}}{1 + \hat{q}_1} < \hat{\theta}_1 < \bar{n} .$$

The ML estimates for the data of Table 1, as obtained by computer, are $\hat{p} = .81$, $\hat{\theta} = 1.02$ and $\hat{\lambda} = 1.53$. This value of p conformed to the expectations of the observers since it was felt that that sampling was fairly intensive. Also, as suspected, $\hat{\lambda} > \bar{k}$ and $\hat{\theta} < \bar{n}$. Average group size, in this example, is estimated simply as $\hat{\theta}g^{(1)}(\hat{\theta})/g(\hat{\theta}) + 1 = \hat{\theta} + 1 = 2.02$.

Estimates of $\sigma_{\hat{\theta}}^2$, $\sigma_{\hat{p}}^2$ and $\text{Cov}(\hat{\theta}, \hat{p})$ were obtained by substituting \hat{p} and $\hat{\theta}$ into U (expression 3.6) and inverting the resulting matrix. When $\hat{\theta}\hat{q}$ is small, as is the case in this example, the series in u_{22} is easily evaluated. The quantities needed are $\gamma_1(x) = x$, $\gamma_2(x) = x^2$, $g(x) = e^x$ and $x = \hat{\theta}\hat{q}^{1+2}$. Since \hat{q} is small, e^x may be approximated by $1 + x$ and, thus, u_{22} obtained as an expression involving simple power series each of which can be evaluated explicitly. This procedure resulted in the estimates:

$$\hat{\sigma}_{\hat{\theta}}^2 = \frac{v_{11}}{42} = .0271$$

$$\hat{\sigma}_{\hat{p}}^2 = \frac{v_{22}}{42} = .0172$$

$$\hat{\text{Cov}}(\hat{\theta}, \hat{p}) = \frac{v_{12}}{42} = -.0072 .$$

$\hat{\sigma}_{\hat{\lambda}}^2$ was estimated by substituting v_{ij} , $\hat{\theta}$ and \hat{p} into (4.4); this produced $\hat{\sigma}_{\hat{\lambda}}^2 = .1327$.

5. Remarks.

The model described here is not limited to aerial census application. The sampling and estimation procedures may be useful whenever objects occur in distinct groups within sampling units, and the size of a group influences its visibility or noticeability. Such situations arise in the laboratory or greenhouse as well as in the field, whenever the thoroughness required for $p = 1$ may be grossly impractical or unattainable. However, it does follow both intuitively and from the properties of the estimators that p should be as close to 1 as possible.

The example presented presumed a Poisson distribution of group size, but other well known distributions such as negative binomial or log are also available. Clearly, certain choices of the power series distribution can yield relatively easier computation of estimates. Also, it should be pointed out

that the authors are not asserting that the distribution of e , in the case of moose, is Poisson. This distribution was chosen for the example because, at least superficially, the choice seems reasonable. In any application, the distribution of e should be carefully chosen, preferably with some biological basis.

Potential users of this model are aware that quadrat size must be chosen so that crowding is not a factor, but should regard themselves admonished to adhere strictly to the sampling procedure.

REFERENCES

- [1] Bergerud, A. T. and Manuel, F. (1969). Aerial census of moose in central Newfoundland. Journal of Wildlife Management 33 910-16.
- [2] Evans, C. D., Troyer, W. A. and Lensink, C. J. (1966). Aerial census of moose by quadrat sampling units. Journal of Wildlife Management 30 767-776.
- [3] Roy, J. and Mitra, S. K. (1957). Unbiased minimum variance estimation in a class of discrete distributions. Sankhyā 18 371-378.